EVIDENCE BASED PUBLIC HEALTH POLICY AND PRACTICE

Effectiveness of a physician-oriented feedback intervention on inappropriate hospital stays

Pedro Antón, Salvador Peiró, Jesús M Aranaz, Rafael Calpena, Antonio Compañ, Edith Leutscher, Vicenta Ruíz

J Epidemiol Community Health 2007;**61**:128–134. doi: 10.1136/jech.2005.040428

See end of article for authors' affiliations

Correspondence to: Dr S Peiró, Fundación Instituto de Investigación en Servicios de Salud, San Vicente 112, 3, 46007 Valencia, Spain; speiro@comv.es

Accepted 13 June 2006

Objective: To evaluate the effectiveness of a combined intervention to reduce inappropriate hospital stays.
Design: Quasi-experimental pre-test/post-test with a non-equivalent control group.
Setting: Three teaching hospitals in the National Health System in Alicante, Spain.
Study participants: Intervention group (2 Surgical Units with 1451 hospital stays) and control group (1 Surgical Unit with 1268 hospital stays).
Intervention: Structured oral presentation followed by direct feedback to surgeons about their own percentages of inappropriate stays and daily evaluation of appropriateness by the surgeons during their rounds.
Main outcome measures: Reduction in the percentage of inappropriate stays identified by the Appropriateness Evaluation Protocol during the intervention period compared to the basal period.
Results: The intervention group reduced its percentage of inappropriate stays from 14.3% to 7.9% (absolute reduction: −6.40; 95% CI −10.7 to −2.14; relative reduction: 44.8%), while no changes occurred in the control group. The reduction was in the number of inappropriate stays attributable to the patients' medical management that went from 12.7% to 5.8% (absolute reduction: −6.92; 95% CI −10.90 to −2.92), while no significant changes occurred in inappropriate stays due to other causes.
Conclusions: A combined intervention of feedback and physician participation in appropriateness evaluations is effective in reducing the percentage of inappropriate hospital stays, particularly those attributable to conservative medical patterns at discharge.

Inappropriate hospital use is defined as days a patient is hospitalised to receive care that, from the clinical perspective, could be provided on a less complex level.[1] Specific instruments such as the Appropriateness Evaluation Protocol (AEP),[2 3] the Oxford Bed Study Instrument[4 5] or the Intensity-Severity-Discharge Criteria Set[6] are used to identify inappropriate hospital use by reviewing the care provided during each hospital stay. Although studies published on this subject have not all used the same methodology, reviews of findings in the US,[7] Europe[8 9] and Spain[10 11] suggest that one out of every three or four hospital stays is inappropriate.

As these striking figures warrant, numerous studies have examined inappropriate use (a search conducted by some of the authors identified 84 papers on the subject in Spain alone through 2002),[12] but very little is known about the effect of interventions in reducing inappropriate stays (the above-mentioned search[12] only identified nine controlled studies published worldwide on the subject).[13–21] Additionally, while the importance of involving professionals in all strategies to improve services has been repeatedly stressed,[22 23] only two of these studies incorporated any mechanism for the doctors themselves to evaluate appropriateness.[16 20] The aim of this study was to evaluate the effectiveness of a mixed intervention (including educational components, feedback and evaluation on the part of the doctors) on the reduction of inappropriate hospital stays in general surgery, using a pre-test/post-test design with a control group. Our research hypothesis was that the group exposed to the intervention would show a decrease in the number of inappropriate stays compared to a basal period prior to the intervention, while the control group would not register any variations.

MATERIAL AND METHODS

Design

A quasi-experimental pre-test/post-test design with a non-equivalent control group.

Setting

The study was conducted in three hospitals in the public healthcare system in Alicante, Spain, that serve 253 000, 229 000 and 196 000 inhabitants respectively. These hospitals are all part of the network of public healthcare centres administered by the Valencia Regional Government. Among the relevant features of the Spanish public healthcare system are the universal and compulsory nature of coverage, the organisation of services by hospital districts, the absence of co-payment schemes except for outpatient pharmacy, and the civil servant-like status of the doctors, who are paid by salary. The three hospitals are all teaching hospitals for under- and postgraduate students, including specialists in general surgery. All have day surgery units and hospital-at-home facilities.

Doctors' groups and study phases

The intervention group comprised the surgeons in the general surgery departments of two of the hospitals, with 53 and 28 beds, and 13 and 9 surgeons, respectively. The average Diagnosis-Related Group (DRG) weight was 1.16 and 0.98, and the bed occupancy rate approximately 80%. The control group was the general surgery department of the third hospital, with 47 beds, 22 surgeons, a DRG weight of 1.18 and a bed

Abbreviations: AEP, Appropriateness Evaluation Protocol; DRG, Diagnosis-Related Group.

occupancy rate of approximately 95% (data referring to the year before the intervention). The study's field work was done between May and June 2000, in three different phases: a basal period of one week (May 8–14), the intervention period of two weeks (May 22–June 4) and, after a two-week hiatus, the post-intervention period (June 19–25).

Intervention

The intervention had various components: educational sessions, feedback to the physicians about their own inappropriate rates and an evaluation of inappropriate stays by the doctors themselves. The message of the feedback exercise[22] was aimed at reducing inappropriate patient-days. The ''messenger'' was one of the researchers, who was not affiliated with any of the three hospitals, and who worked with the explicit support of the heads of both of the participating surgery departments. The ''recipients'' of the message were the surgeons in the intervention group. The intervention took place over a two-week period, with the format of group sessions with oral presentations and discussions and a review of written reports and scientific articles on the topic. After receiving this information, the doctors were asked to evaluate the appropriateness of their own patients' hospital stays by using an instrument designed specifically for this study called the adeQhos[24] during their daily rounds. The intervention was sequenced as follows:

1. A structured presentation was given to the department heads of the three participating hospitals to explain the concept of inappropriate hospitalisation and its most frequent causes, followed by a discussion of data published by other hospitals gathered from a review of the literature; during these sessions the support of the department heads for the project was obtained.

2. Data from the participating departments were collected, done simultaneously in the three hospitals during every phase of the study.

3. A structured presentation was given to the doctors in the intervention group to explain the different aspects of inappropriate hospital use, the type of intervention to be implemented and the mechanisms of the self-evaluation component.

4. Daily evaluation was made of the appropriateness of hospital stays by the surgeons in the intervention group using the adeQhos, for two consecutive weeks.

5. Feedback was provided to the doctors in the intervention group about their own inappropriate hospitalisation rates measured during the basal period.

The intervention did not include any incentives, financial or otherwise, nor was the hospital's Administration involved, although the support of the heads of the respective surgery departments lent authority to the project.

The instrument used for the doctors in the intervention group to evaluate the appropriateness of their own patients' stays was the adeQhos. It was designed for physicians to use at the patient's bedside; its structure is analogous to the AEP and the Oxford Bed Study Instrument. The instrument's clinimetric properties and the process followed for its validation are described elsewhere.[24] The adeQhos has four headings: (1) data to identify the day under review; (2) criteria to evaluate the appropriateness of the hospital day; (3) a subjective evaluation of the appropriateness of the hospital day by the physician; and (4) a list of causes of inappropriate hospital stays (table 1). The questionnaire is produced in a self-adhesive ''post-it'' format for easy use and reference. The use of the adeQhos during the intervention was as follows:

1. The research team recorded data about the day selected for review before the doctors' daily round.

2. The adeQhos was ''posted'' in a visible place on the medical record.

3. Each attending physician answered yes/no questions about the patient to establish the appropriateness of the hospitalisation day during his round; for a hospitalisation day to be considered appropriate, it was necessary for the doctor to mark any one of the criteria on the list, and in these cases the review was concluded.

4. If the day did not meet any of the criteria, the doctor needed to use his clinical judgment and decide on the basis of medical criteria alone whether the patient needed to continue to be hospitalised; if the decision was ''yes'', the day was considered appropriate from the clinical perspective and the evaluation was concluded.

5. If the answer was ''no'', the doctor had to select one of the reasons for an inappropriate day from the list provided. The surgeons in the intervention group completed the adeQhos daily during the intervention period as they examined their patients during their daily rounds. The median time required to fill in the form was one minute.

Outcomes measured

The main outcome measured was the reduction in the percentage of inappropriate hospital stays, identified with the AEP,[2 3] during the intervention period compared to the basal period. Additionally, since the intervention had been targeted to the doctors and was not intended to change aspects of hospital organisation, secondary outcomes measured were the reduction in inappropriate stays attributable to ''medical discharge management'' and ''programming problems'', as defined on the list of reasons for inappropriate use in the AEP.

The version of the AEP that was used in this study was one designed for adult medical-surgical patients that had been previously validated in Spain[25] and complemented with the Spanish version of the AEP Manual.[26] The AEP contains two sets of criteria that are objective and independent of the diagnosis, designed to identify inappropriate admissions and inappropriate stays respectively, from the medical perspective alone (excluding social, family-related and all other non-medical topics) in non-psychiatric and non-obstetric adult patients. The criteria evaluating the stay include 26 items referring to medical and nursing care and the patient's clinical condition that must be reviewed for each hospital day. If any one of these criteria are met on any given day, the stay is considered to be appropriate. The AEP also includes a list of causes of inappropriate hospital days that for this study were grouped under four headings: (1) problems associated with programming the operating theatre, diagnostics tests, pre-surgery preparation, etc, during which time the patient must be hospitalised (reasons 1–7 in the original version),[26] (2) problems attributable to conservative medical management patterns, primarily those affecting discharge (reason 8 in the original version),[26] (3) problems affecting discharge but attributable to the patient or the family (reason 9 in the original version),[26] and (4) problems related to discharge attributable to deficiencies in healthcare alternatives other than hospitalisation (reason 10 in the original version).[26]

Sample

The unit of analysis was the days the patients were hospitalised in the participating general surgery departments during the three periods comprising the study. The study excluded days corresponding to children under the age of 8 (which need to be evaluated with the paediatric AEP), the day of admission (evaluated on the AEP with the Admissions Set), the day of discharge (not evaluated according to the AEP Manual) and the days the patients were hospitalised in other departments not participating in the intervention (intensive care, oncology or other medical departments). The necessary sample size was

Table 1 Description of the instrument used for the doctors to evaluate the appropriateness of their own patients' stays (adeQhos[24])

Answer ''yes'' or ''no'' for each question in relation to this patient	1	Is any intervention going to take place TODAY that required the patient to be hospitalised all day yesterday for evaluation or pre-op preparation?
	2	Did the patient receive surgery, cardiac catherisation, angiography, biopsy of any internal organ, thoracocentesis, paracentesis or invasive procedures of the central nervous system YESTERDAY or THE DAY BEFORE YESTERDAY?
	3	Did the patient experience incapacity to urinate and/or defecate (not attributable to neurological problems or chronic constipation) YESTERDAY or THE DAY BEFORE YESTERDAY?
	4	Were any tests performed YESTERDAY that required strict diet control? Was the patient monitored by a doctor (>3 visits/day)? Was the patient taking any new (or experimental) treatment requiring frequent dose adjustments under direct medical supervision?
	5	Did the patient receive respiratory therapy and/or mechanical ventilation through inhalation (at least three times a day) YESTERDAY? Did s/he receive any form of parenteral therapy? i.m. or s.c. injections (at least 2 times per day, excluding insulin)?
	6	YESTERDAY, were the patient's vital sign monitored (at least every 30 minutes during at least 4 hours)? Was his/her water equilibrium corrected with medication? Did the patient require monitoring (at least 3 times/day under medical supervision)? Did the patient receive treatment for surgical wounds and/or draining?
	7	YESTERDAY or THE DAY BEFORE YESTERDAY did the patient receive any transfusions? Did s/he present ventricular fibrilation or was acute ischaemia evidenced with ECG? Did the patient have a fever (>38° rectal) that was not related to the cause of admission? Was the patient in a coma or with acute confusion? Were there signs of acute haematological alterations? Did the patient present progressive neurological difficulties?
If you answered ''no'' to all of the questions, it is likely that from a strictly medical standpoint, and without taking into consideration problems related to the family or organisational logistics, etc, this patient would not have needed to remain hospitalised yesterday. Do you believe that for CLINICAL REASONS this patient needs to remain hospitalised? (*yes/no*)		
If you answered ''no'' to the previous question, the patient's stay is due to (check reason):	1	Problems of programing diagnostic tests
	2	Problems of programing surgery
	3	Premature admission
	4	Problems with the patient or family
	5	Responsibility of the hospital: organisational problems with discharge
	6	Lack of alternative care (overnight facilities, hospital-at-home, etc)
	7	Responsibility of the doctor: excessively conservative length of stay as part of patient management
	8	Others (specify):

estimated to be 159 (basal period) and 318 days (intervention period) in order to detect a reduction of 25% to 15% (absolute reduction of around 10% and relative reduction of 40%, one-sided) between the two periods in the percentage of inappropriate stays, with an alpha error of 0.05, a power of 0.80 and a ratio of 1:2 (the days of one week during the basal period and two weeks during the intervention period). The percentage of inappropriate stays during the basal period was calculated after reviewing Spanish studies on appropriateness in general surgery. The reduction of 15–25% was selected in order to obtain a rate that would be relevant in the context of the setting where the field work was to be conducted. The sample ultimately obtained—all hospital stays during the three periods of the study—was larger than the minimum size calculated (n = 2719 hospital stays; control group: 1268, basal: 274, intervention: 648, post-intervention: 346; intervention group: 1451, basal: 322, intervention: 774, post-intervention: 365).

Study process

The medical records of all the patients who met inclusion criteria during the study's different periods, including weekends, were reviewed by two doctors and a nurse who had received previous training using records not included in the study. After studying the AEP Manual,[26] the reviewers evaluated 40 hospital stays independently. The usual concordance indices between reviewers were then analysed,[27] giving an overall agreement index of 0.97 and 1.00, a specific agreement index of 0.80 and 1.00 and kappa statistics of 0.87 and 1.00, and thus showing excellent agreement among the reviewers. The medical records were reviewed on the wards

Table 2 Admissions characteristics of the intervention and control groups during basal, intervention and post-intervention periods (percentages in parentheses)

		Control group			Intervention group		
		Basal	Intervention	Post-intervention	Basal	Intervention	Post-intervention
Age (years)	<45	14 (42.4)	25 (34.2)	5 (22.7)	23 (35.4)	38 (25.0)	22 (35.5)
	45–64	7 (21.2)	27 (37.0)	6 (27.3)	16 (24.6)	59 (38.8)	18 (29.0)
	>64	12 (36.4)	21 (28.8)	11 (50.0)	26 (40.0)	55 (36.1)	22 (35.5)
Sex	Female	17 (51.5)	28 (38.4)	10 (45.6)	26 (40.0)	69 (45.4)	24 (38.7)
	Male	16 (48.5)	45 (61.6)	12 (54.5)	39 (60.0)	83 (54.6)	38 (61.3)
Days*	Weekdays	26 (78.8)	56 (76.7)	22 (100.0)	50 (76.9)	123 (80.9)	50 (80.6)
	Weekend	7 (21.2)	17 (23.3)	0 (0.0)	15 (23.1)	29 (19.1)	12 (19.4)
Surgery	No	4 (12.1)	11 (15.1)	7 (31.8)	11 (16.9)	32 (21.1)	14 (22.6)
	Yes	29 (87.9)	62 (84.9)	15 (68.2)	54 (83.1)	120 (78.9)	48 (77.4)
Admission*	Programmed	21 (63.6)	33 (45.2)	6 (27.3)	37 (56.9)	80 (52.6)	29 (46.8)
	Emergency	12 (36.4)	40 (54.8)	16 (72.7)	28 (42.1)	72 (47.4)	33 (53.2)
LOS	<7 days	20 (60.6)	41 (6.2)	11 (50.0)	45 (69.2)	111 (73.0)	37 (59.7)
	7+ days	13 (39.4)	32 (43.8)	11 (50.0)	20 (30.8)	41 (27.0)	25 (40.3)
Total	Admissions	33	73	22	65	152	62
	Days	274	648	346	322	774	355

LOS, length of stay.
$p<0.05$ in the χ^2 or Fisher's test.

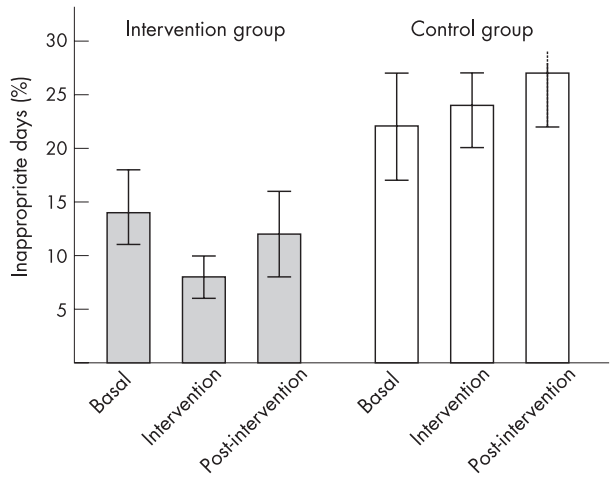

Figure 1 Percentage of inappropriate days by study periods in the intervention and control groups.

while the patients were hospitalised. All of the documents contained in these records were available and used in the reviews, including medical orders, nursing notes and graphs, interconsultation reports, diagnostic tests, therapeutic measures and the results of laboratory tests.

Authorisations, confidentiality and ethics of the research

Before implementing the project, authorisation was obtained from the three hospitals' administrations. The heads of the general surgery departments of the three participating hospitals also lent their cooperation. The research was further approved by the research and ethics committees of the hospitals where the intervention took place. Neither the adeQhos nor the database used incorporated any information identifying the patients, except for the number of clinical record (preserved in case a record needed to be retrieved for review). This number was retained in a database separate from the one used for the analysis of the data, to which only the head researcher had access.

Statistical analysis

A descriptive analysis of the characteristics of the patients whose hospital days were analysed was performed with the χ^2 test to check for possible changes between groups and periods. When appropriate, Fisher's exact test was used instead of the χ^2. Next, the effectiveness of the intervention was evaluated, by comparing the percentage of inappropriate stays calculated for the basal and intervention periods in the intervention group, using the test for differences between proportions. Because the study's design revolved around a pre-established hypothesis, one-sided tests were used to evaluate a single outcome, the reduction of the percentage of inappropriate stays. Additionally, the differences between the basal and post-intervention periods, and intervention and post-intervention periods were evaluated. The same analyses were performed in the CG in order to test the complementary hypothesis of the absence of changes in this group. Next, the various causes of inappropriate stays were analysed, along with the impact of the intervention on the secondary outcomes (inappropriate stays due to medical management and programming problems). In all cases, the corresponding confidence intervals of 95% (95% CI) are shown for the percentages and differences in the values calculated using the exact binomial method. All the statistical analyses were done with the SPSS (SPSS Inc, Chicago, Illinois, USA) and Stata (Stata Corporation, College Station, Texas, USA) statistical programs.

RESULTS

Table 2 shows the characteristics of the admissions with hospital days that were evaluated in both groups during the three periods of the study. No differences were found between the groups for any of the periods in terms of age, sex, percentage of patients receiving surgery or the percentage of patients with hospital stays exceeding one week. Approximately 20% of the stays occurred during the weekend, with a homogeneous distribution, with the exception of the control group during the post-intervention period, where there were no admissions during the weekend. In the control group, the percentage of emergency admissions rose throughout the study.

The percentage of inappropriate stays for the whole study was 16.7% (95% CI 15.3 to 18.2). Table 3 and figure 1 show the overall percentage of inappropriate stays during each of the study's periods. The control group did not show any statistically significant differences in the percentage of inappropriate stays between the periods (21.5%, 23.6% and 27.2%). In the intervention group, there was a reduction in inappropriate stays from 14.3% in the basal period to 7.9% during the intervention period (representing an absolute reduction of 6.4 points, and a relative reduction of 44.8%), with a rise to 11.8% in the post-intervention period. Both hospitals whose surgery departments comprised the intervention group showed similar behaviour throughout the study with inappropriate stays rates of 17.5%, 8.5% and 13.1% in the first hospital, and 12.5%, 7.5% and 11.05 in the other, for the basal, intervention and post-intervention periods, respectively.

For the three study periods overall, the breakdown by causes of inappropriate use showed that programming problems accounted for the greatest number of inappropriate stays in the control group (51.0% of the inappropriate stays) followed by medical management problems (45.7%). In the intervention group, however, programming problems only accounted for 10.0% of the inappropriate stays, while medical management problems accounted for 79.2%. Problems involving the patient or family, or deficient alternative resources had minimum

Table 3 Percentage of inappropriate stays (all reasons) and differences between periods in the intervention and control groups

		Control group		Intervention group	
		n (%)	95% CI	n (%)	95% CI
Period	Basal	274 (21.53)	16.81 to 26.87	322 (14.29)	10.65 to 18.59
	Intervention	648 (23.61)	20.39 to 27.07	774 (7.88)	6.08 to 10.01
	Post-intervention	346 (27.17)	22.54 to 32.18	355 (11.83)	8.66 to 15.65
	Total	1268 (24.13)	21.80 to 26.58	1451 (10.26)	8.75 to 11.94
Differences	Basal *v* intervention	2.08	−3.79 to 7.94	−6.40	−10.7 to −2.14

Table 4 Percentage of inappropriate stays attributable to medical management and differences between periods in the intervention and control groups

		Control group		Intervention group	
		n (%)	95% CI	n (%)	95% CI
Inappropriate stays by periods	Basal	274 (11.31)	7.81 to 15.67	322 (12.73)	9.29 to 16.87
	Intervention	648 (9.56)	7.41 to 12.09	774 (5.81)	4.27 to 7.70
	Post-intervention	346 (13.58)	10.15 to 17.65	355 (9.01)	6.24 to 12.48
	Total	1268 (11.04)	9.36 to 12.89	1451 (8.13)	6.77 to 9.65
Differences	Basal *v* intervention	−1.75	−6.13 to 2.64	−6.92	−10.90 to −2.92

effects in both groups. A breakdown of the proportion of inappropriate stays attributable to medical management problems (table 4) showed behaviour similar to inappropriate use for all days: there were no significant differences in the control group (11.3%, 9.6% and 13.6% in the respective periods), while in the intervention group, there was a statistically significant reduction in inappropriate stays due to this cause, going from 12.7% in the basal period, to 5.8% during the intervention period, with a rise to 11.0% during the post-intervention period. The breakdown of inappropriate stays due to programming problems (table 5) revealed that neither group showed significant changes during the study's periods.

DISCUSSION

The results of this study show that in the intervention group there was a marked drop in the percentage of inappropriate stays after the intervention, while in the control group no significant changes occurred. These findings support the hypothesis of the effectiveness of the intervention. Furthermore, the reduction occurred in the inappropriate stays attributable to medical management problems (from 12.7% to 5.8%), indicating that the intervention affected the situations that were directly under the doctors' control, confirming the success of the exercise. The literature contains very few controlled studies that have evaluated the impact of interventions on inappropriate hospital use,[13–21] and most of those published refer to the context in the US[13–15 17] or Israel,[16 20] environments that are very different from the Spanish National Health System. None of the studies published has concentrated exclusively on a hospital's surgery department. With the exception of two,[13 9] all used the AEP, and most were based on quasi-experimental designs,[14 16–21] with interventions with feedback components.[13 14 16–19] Generally, all of the studies show positive results in terms of the reduction of the percentage of inappropriate stays, although in a few cases these are not conclusive due to the small sample size.[18] The only two studies where the doctors conducted their own evaluations (both from Israel) show results that are very similar to those in this study, with reductions of 52.6% of inappropriate stays in medical services[16] and 54.6% in paediatric services.[20] The studies evaluating feedback-based interventions without control groups also showed positive results[28–31] that are, in general, consistent with our own.

Certain precautions must be taken when evaluating the effectiveness of an intervention. The internal validity of any study with a before/after design and a non-equivalent control group may be threatened. However, it is not likely that the effect detected here was due to other interventions (administrative or management changes, etc) rather than the intervention evaluated, because the intervention took place over a short period of time. The three hospitals involved belong to the same healthcare organisation, and no distinction was detected in the strategies deployed by the hospitals during the course of the study. It is also unlikely that a regression to the mean or seasonal variations would explain the changes registered in the intervention group. Any such effect would be expected to be greater in the control group, which had higher basal levels. Further, it is not likely that the results obtained were due to differences between the control group and the intervention group rather than to the intervention. Although it is true that, because of the study's design, there might have been differences between the groups, the types of patients receiving care were relatively similar between the groups and periods (table 2). The higher percentage of inappropriate stays registered in the control group during the basal period is due to programming problems, while problems associated with medical management issues—the main object of the intervention—were similar in both groups. What is more, the levels of inappropriate stays reached in the intervention group during the intervention period, among the lowest reported in the literature, would be difficult to achieve by chance, and the two hospitals whose surgery departments comprised this group showed similar behaviour throughout the study—factors that reinforce the causal relation between the intervention and the outcomes.

The possibility of contamination during the study is remote, although the doctors in the control group were probably aware that the study was taking place (the head of the surgery department of the control group was informed of the study, the hospitals are all relatively close to each other and the doctors

Table 5 Percentage of inappropriate stays attributable to programming problems and differences between periods in the intervention and control groups

		Control group		Intervention group	
		n (%)	95% CI	n (%)	95% CI
Inappropriate stays by periods	Basal	274 (10.21)	6.90 to 14.43	322 (0.31)	0.00 to 1.72
	Intervention	648 (14.0)	11.46 to 16.96	774 (0.90)	0.36 to 1.85
	Post-intervention	346 (10.7)	7.64 to 14.44	355 (1.97)	0.80 to 4.02
	Total	1268		1451	
Differences	Basal *v* intervention	3.82	−0.65 to 8.30	0.59	−0.30 to 1.50

likely to know each other). However, any contamination would have had the effect of prompting the control group to equal the intervention group's behaviour, and would not explain the reductions in inappropriate stays in the latter. Finally, the Hawthorne effect and other similar effects (the effect of feeling oneself observed, of novelty and social desirability) may have contributed to some extent to the results obtained, particularly with the involvement of the department heads and the high visibility of the research with the use of review sheets on the wards. Nevertheless, these effects were built into the intervention. Results might have been different without the support of the department heads or had a less visible research method been used.

There are other aspects that impose limitations on a study of this nature. The limitations of the AEP itself when identifying inappropriate admissions and days, extensively discussed elsewhere,[32–34] would affect all the groups in the same way. A lack of fit due to the severity of the cases, in spite of the higher DRG weight of the control group, is not likely to have influenced the results, since the greater severity and intensity of care in the control group (given the sensitivity of the AEP to the severity of the patients and the intensity of care provided) should have translated into a lower percentage of inappropriate use. The quality of the medical records would also be expected to affect appropriateness rates (records with less documentation would produce a higher inappropriateness rate than the actual one, because they would not be able to justify the appropriateness of a day),[34] but this aspect weakens with an intragroup before/after analysis (rather than a comparison between groups). Furthermore, this aspect could only have a discrete effect at best, as an analysis of the degree of association between the quality of the clinical records and the percentage of inappropriate stays showed there was no correlation between them.[33] Finally, it should be pointed out that in this analysis, as usually occurs in utilisation reviews, related events (the occasional successive days of the same patient) were treated as independent events. This aspect presents both drawbacks and advantages in terms of the use of more complex statistical methods.[35]

When evaluating the impact of an intervention, it is difficult to isolate the direct impact from aspects associated with its context (from the way the problem is focused, to the procedures followed, methods used to ensure collaboration, the development of easy-to-use instruments, the empathy of the evaluators and the organisation's ''climate''). What is more, surely no intervention alone is capable of magically increasing the quality and efficiency of healthcare services anywhere.[36] Nonetheless, a wide range of interventions are at our disposal to achieve substantial improvements in the effectiveness and efficiency of healthcare services, and evaluating the measures implemented makes it possible to discern which are best for each context. The results of this study, consistent with those published elsewhere,[16 20] suggest that combined interventions are effective in achieving major reductions in the percentage of inappropriate stays, especially in stays attributable to medical management issues. The elements key to the effectiveness of these interventions are most likely the use of active feedback (known to be more effective than merely communicating passive information[37]), reinforcement by authority (in this case through the support of the heads of the surgery departments), and the direct involvement of the doctors who evaluated the appropriateness of their patients' hospitalisation days with the easy-to-use questionnaire filled in on the ward.

What is already known

- Several studies in the US and Europe suggest that inappropriate hospital use (hospital stays that could be provided on a less complex level as outpatient care, primary care or others) accounts for one out of every three or four hospital stays.
- Very little is known, in particular in public healthcare systems, about the effectiveness of interventions aimed to reduce inappropriate stays.

What this paper adds

- A mixed intervention (with educational sessions, feedback and evaluation conducted by the doctors themselves) is effective in reducing inappropriate stays. The reduction primarily affects inappropriate stays attributable to medical management problems.

Policy implications

- Strategies involving the direct participation of doctors can be useful in reducing inappropriate hospitalisation days.
- These strategies are effective in public healthcare systems with scarce economic incentives.

Authors' affiliations

P Antón, S Peiró, Fundación Instituto de Investigación en Servicios de Salud, Valencia, Spain
P Antón, V Ruíz, Inspección de Servicios Sanitarios, Elche, Spain
J M Aranaz, Servicio de Medicina Preventiva, Hospital Universitario de San Juan, San Juan and Departamento de Salud Pública, Universidad Miguel Hernández, Elche, Spain
R Calpena, Servicio Cirugía General, Hospital Universitario de Elche, Elche, Spain
A Compañ, Servicio Cirugía General, Hospital Universitario de San Juan, San Juan, Spain
E Leutscher, Servicio de Medicina Preventiva, Hospital de Villajoyosa, Villajoyosa, Spain

This work is part of a research project financed by the Fundación Mapfre Medicina (Mapfre Medical Foundation) (1999/2000) and the Fundación Instituto de Investigación en Servicios de Salud (Institute for Health Services Research Foundation).

REFERENCES

1 **Peiró S**, Meneu R. Utilization review. Definition, concept, methods [in Spanish]. *Rev Calidad Asistencial* 1997;**12**:122–36.
2 **Gertman PM**, Restuccia JD. The Appropriateness Evaluation Protocol: a technique for assessing unnecessary days of hospital care. *Med Care* 1981;**19**:855–71.
3 **Restuccia JD**, Payne SMC, Lenhart GM, *et al*. Assessing the appropriateness of hospital utilization to improve efficiency and competitive position. *Health Care Manag Rev* 1987;**12**:17–22.
4 **Anderson P**, Meara J, Brodhurst S, *et al*. Use of hospital beds: a cohort study of admissions to a provincial teaching hospital. *BMJ* 1988;**297**:910–2.
5 **Victor CR**, Nazareth B, Hudson M. The inappropriate use of acute beds in an inner London District Health Authority. *Health Trends* 1993;**25**:94–7.
6 **Inglis AL**, Coast J, Gray SF, *et al*. Appropriateness of hospital utilization. The validity and reliability of the Intensity-Severity-Discharge Review System in a United Kingdom acute hospital setting. *Med Care* 1995;**33**:952–7.
7 **Peiró S**, Portella E. Identification of the inappropriate use of hospitalization: the search for efficiency [in Spanish]. *Med Clin (Barc)* 1994;**103**:65–71.

8 **Liberatti A**, Apolone G, Lang T, *et al*. A European project assessing the appropriateness of hospital utilization: background, objectives and preliminary results. *Int J Qual Health Care* 1995;**7**:187–99.
9 **Lorenzo S**, Beech R, Lang T, *et al*. An experience of utilization review in Europe: sequel to a BIOMED proyect. *Int J Qual Health Care* 1999;**11**:13–19.
10 **Lorenzo S**, Suñol R. An overview of Spanish studies on appropriateness of hospital use. *Int J Qual Health Care* 1995;**7**:213–18.
11 **Meneu R**, Peiró S. Review of the inappropriate hospital use in Spain: from scientifical communication to practical utilization? [in Spanish]. *TodoHosp* 1997;**134**:53–60.
12 **Antón García P**. *Impact of an intervention on inappropriate hospitalization in the general surgery services of the Alicante Universitary Hospitals* (Thesis) [In Spanish]. Elche: Departamento de Salud Pública, Universidad Miguel Hernandez, 2003.
13 **Restuccia JD**. The effect of concurrent feedback in reducing inappropriate hospital utilization. *Med Care* 1982;**20**:46–62.
14 **Restuccia JD**, Payne SMC, Welge CH. *Reducing inappropriate use of inpatient medical/surgical and pediatric services*. Baltimore: Health Care Financing Administration, 1986.
15 **Siu AL**, Sonnenberg FA, Manning WG, *et al*. Inappropriate use of hospitals in a randomized trial of health insurance plans. *N Engl J Med* 1986;**315**:1259–66.
16 **Mozes B**, Halkin H, Katz A, Schiff E. Reduction of redundant hospital stay through controlled intervention. *Lancet* 1987;**1**:968–9.
17 **Payne SMC**, Ash A, Restuccia JD. The role of feedback in reducing medically unnecessary hospital use. *Med Care* 1991;**29**(Suppl):AS91–106.
18 **Spolaore P**, Tasso S, Lustro L, *et al*. Valutazione della appropriatezza di ricoveri di pazienti anziani: analisi delle variazioni dopo attivazione dei servizi di assistenza infermieristica domiciliare [in Italian]. Atti V Congreso Nazionale Società Italiana di VRQ. Padova, 1992.
19 **Cardiff K**, Anderson G, Sep S. Evaluation of a hospital-based utilization management program. *Health Manage Forum* 1995;**8**:38–45.
20 **Vardi A**, Modan B, Blumstein Z, *et al*. A controlled intervention in reduction of redundant hospital days. *Int J Epidemiol* 1996;**25**:604–8.
21 **Moya-Ruiz C**, Peiró S, Meneu R. Effectiveness of feedback to physicians in reducing inappropriate use of hospitalization. *Int J Qual Health Care* 2002;**14**:305–12.
22 **O'Brien MAT**, Oxman AD, David DA, *et al*. Audit and feedback: effects on professional practice and health care outcomes (Cochrane Review). In: *The Cochrane Library, Issue 3, 2001*. Oxford: Update Software).
23 **Restuccia JD**. Strengths and Weaknesses of Hospital Appropriateness Tools. 2nd Convegno Nazionale PRUO. Bari, Italy: May, 2002:15–18.
24 **Antón P**, Peiró S, Aranaz JM, *et al*. The adeQhos, a tool to gauge appropriateness of hospitalization during the ward round [in Spanish]. *Rev Calidad Asistencial* 2002;**17**:591–9.
25 **Peiró S**, Meneu R, Roselló ML, *et al*. Validity of the protocol for evaluating the inappropriate use of hospitalization [in Spanish]. *Med Clin (Barc)* 1996;**107**:124–9.
26 **Restuccia JD**. Appropriateness Evaluation Protocol. Manual [in Spanish]. Barcelona: Fundación Avedis Donabedian, 1995, (adapted from:Restuccia JD.Appropriateness Evaluation Protocol Reviewer's Manual.In:*The AEP Final Report*.Publication PB87–112041.Springfield, VA:National Technical Information Service, 1986)..
27 **Fleiss JL**. The measurement of interrater agreement. In:Fleiss JL (ed). *Statistical methods for rates and proportions*. New York: John Wiley & Sons, 1981:212–36.
28 **Studnicki J**, Stevens CE, Knisely L. Impact of a cybernetic system of feedback to physicians on inappropriate hospital use. *J Med Educ* 1985;**60**:454–60.
29 **Restuccia JD**, Payne SMC, Lenhart GM, *et al*. Assessing the appropriateness of hospital utilization to improve efficiency and competitive position. *Health Care Manag Rev* 1987;**12**:17–22.
30 **Payne SMC**, Restuccia JD, Ash A, *et al*. Using utilization review to improve hospital efficiency. *Hosp Health Serv Admin* 1991;**36**:473–90.
31 **Payne SMC**, Campbell D, Penzías BG, *et al*. New methods for evaluating utilization management programs. *Qual Rev Bull* 1992;**18**:340–7.
32 **Payne SMC**. Identifying and managing inappropriate hospital utilization: a policy synthesis. *Health Serv Res* 1987;**22**:709–69.
33 **Ramos-Cuadra A**, Marión-Buen J, García-Martín M, *et al*. Use of the Appropriateness Evaluation Protocol: the role of medical record quality. The effect of completeness of medical records on the determination of appropriateness of hospital days. *Int J Qual Health Care* 1995;**7**:267–75.
34 **Lorenzo S**. Utilization review methods: limitations [in Spanish]. *Med Clin (Barc)* 1996;**107**:22–5.
35 **Gange SJ**, Muñoz Alvaro, Saez M, et al. Use of the beta-binomial distribution to model the effect of policy changes on appropriateness of hospital stays. *Appl Stat* 1996;**45**:371–82.
36 **Oxman AD**, Thomson MA, Davis DA, *et al*. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *Can Med Assoc J* 1995;**153**:1423–31.
37 **Mugford M**, Banfield P, O'Hanlon M. Effects of feedback of information on clinical practice: a review. *BMJ* 1991;**303**:398–402.

APHORISM

The lessons from BSE

If BSE should have taught us anything, it is that no evidence of proof is not evidence of no proof. Not understanding this was a big part of the problem with the way in which this crisis was mishandled.

JRA